

The symmetry relationships between pairs of connectrons

Reference to Related Applications

- 5 The present application includes the subject of Provisional Application Serial No. 60/455,563 filed March 19,2003

The present application is a continuation in part of US Patent Application Serial Number 09/866,925 filed May 30, 2001
10 entitled ALGORITHMIC DETERMINATION OF FLANKING DNA SEQUENCES THAT CONTROL THE EXPRESSION OF SETS OF GENES IN PROKARYOTIC, ARCHEA AND EUKARYOTIC GENOMES. (referred to as "basic methods patent application")

- 15 This present application is related to PCT application PCT/US01/16471 filed May 31, 2001 and entitled ALGORITHMIC DETERMINATION OF FLANKING DNA SEQUENCES THAT CONTROL THE EXPRESSION OF SETS OF GENES IN PROKARYOTIC, ARCHEA AND EUKARYOTIC GENOMES.

20 The present application is also related to US Patent Application Serial Number 10/339,666 filed January 10, 2003 entitled SIMULATION OF GENE EXPRESSION CONTROL USING CONNECTRONS, INTERFERENCE RNAS (IRNAS) AND SMALL TEMPORAL RNAS
25 (STRNAS) IN PROKARYOTIC, ARCHEA AND EUKARYOTIC GENOMES

The present application is also related to US Patent Application Serial Number 10/364,516 filed February 12, 2003 entitled DETERMINATION OF FLANKING DNA SEQUENCES THAT CONTROL
30 THE EXPRESSION OF SETS OF GENES IN THE ESCHERICHIA COLI K-12 MG1655 COMPLETE GENOME

The present application is also related to US Patent Application Serial Number 10/364,412 filed February 12, 2003 entitled DETERMINATION OF FLANKING DNA SEQUENCES THAT CONTROL THE EXPRESSION OF SETS OF GENES IN THE SACCHAROMYCES CEREVISIAE COMPLETE GENOME

Introduction

The connectron structure of a genome determines sets of four DNA sequences of minimum length of 15-bases (C1 and C2 which are in the 3'UTR of a gene or pseudogene, and T1 and T2 which bracket a set of genes or pseudogenes). The connectrons in a genome control the expression of sets of genes. This patent application describes new types of connectrons as well as how pairs of equivalent and non-equivalent RNA sequences can bind to double-stranded DNA to form a variety of connectrons.

Definitions

Previous definitions of connectron structure are included by reference.

Connectrome - All the connectrons in a given genome.

Dominant Direction - The DNA sequence of a chromosome or a genome from the 5' end to the 3' end of the positive strand.

Anti-Dominant Direction - The DNA sequence of a chromosome or a genome from the 5' end to the 3' end of the negative strand.

C1/C2 Polarity - The direction of the binding of the RNA of a connectron in the major groove of the double-stranded DNA in the dominant direction or the anti-dominant direction.

Uni-Polar C1/C2 - The binding of the RNA of a connectron such that either (1) both the C1 sequence and the C2 sequence bind in the dominant direction or (2) both the C1 sequence and the C2 sequence bind in the anti-dominant direction.

Bi-Polar C1/C2 - The binding of the RNA of a connectron such that either (1) the C1 sequence binds in the dominant direction and the C2 sequence binds in the anti-dominant

direction or (2) the C1 sequence binds in the anti-dominant direction and the C2 sequence binds in the dominant direction.

Reverse Complement - Going away from a given point, the same sequence occurs on opposite strands. In the example below the sequence GCATCC in the dominant direction of the positive strand occurs somewhere else in the genome in the anti-dominant direction of the negative strand

10 Positive Strand 5'- **GCATCC**GTGTAAT ATTACACGGATGC -3'
Negative Strand 3'- CGTAGGCACATTA TAATGTG**CCTACG** -5'

Equivalent Sequences - Two sequences such that the second sequence is in the reverse complement of the first sequence

15 First sequence 5'- GCATCCGTGTAAT -3' (A)

Second sequence 5'- ATTACACGGATGC -3' (A')

20 If the first sequence is called A then the second sequence is called A'

Symmetric Lower-Upper Pair of Connectrons - The binding of two equivalent uni-polar RNA C1/C2 sequence pairs to double-stranded DNA.

Asymmetric Lower-Upper Pair of Connectrons - The binding of two non-equivalent uni-polar RNA C1/C2 sequence pairs to double-stranded DNA.

Symmetric Left-Right Pair of Connectrons - The binding of two equivalent bi-polar RNA C1/C2 sequence pairs to double-stranded DNA.

Asymmetric Left-Right Pair of Connectrons - The binding of two non-equivalent bi-polar RNA C1/C2 sequence pairs to double-stranded DNA.

Connectron Lifetime - A time that varies directly with the length of the shorter of the two triple-stranded generalized Hoogsteen helices formed by the binding of the C1 and C2 RNA connectron sequences to the major groove of the double-stranded DNA.

Connectron Pair Lifetime - A time that varies directly with the product of the lifetimes of the two connectrons in the pair.

Specificity of a Pair of Connectrons - The number of similar or different C1/C2 sources needed to form the pair of connectrons.

Symmetric Connectron Specificity - The specificity of a pair of connectrons formed with equivalent uni-polar or bi-polar RNA sequences.

5

Asymmetric Connectron Specificity - The specificity of a pair of connectrons formed with non-equivalent uni-polar or bi-polar RNA sequences.

10 **Competitive Mode of Behavior in the Formation of a Connectron Pair** - The situation where two different genes produce the same uni-polar or bi-polar C1/C2 sequences of the same or different lengths that bind to the major groove of the double-stranded DNA to form a connectron pair.

15

Cooperative Mode of Behavior in the Formation of a Connectron Pair - The situation where two different genes produce different uni-polar or bi-polar C1/C2 sequences of the same or different lengths that bind to the major groove of the double-stranded DNA to form a connectron pair such that the

20 connectron pair could only be formed from the two different C1/C2 sequences.

25 **Prior Art**

Included by reference.

Brief Description of the Objects of the Invention

5 The basic methods patent application provides the methods for determining the structure of the connectrons in a variety of prokaryotic, Archeal and eukaryotic genomes.

An object of this invention is to provide a method for
10 identifying a one or more new classs of connectrons that bind to the major groove of double-stranded DNA in two directions.

An object of this invention is to provide a method for
15 designing a new class of connectrons that bind to the major groove of double-stranded DNA in two directions.

An object of this invention is to provide a method for
20 identifying the relationship between a pair of connectrons in a genome.

An object of this invention is to provide a method for
designing the relationship between a pair of connectrons in a genome.

25 An object of this invention is to provide a method for identifying the relationship between an existing pair of connectrons in a genome that act in competitive mode such that with respect to the individual connectrons there is an increased lifetime of connectron control of a set of genes.

30 An object of this invention is to provide a method for designing a new synthetic pair of connectrons in a genome that act in competitive mode such that with respect to the

individual connectrons there is an increased lifetime of connectron control of a set of genes.

5 An object of this invention is to provide a method for identifying the relationship between an existing pair of connectrons in a genome that act in cooperative mode such that with respect to the individual connectrons there is an increased lifetime of connectron control of a set of genes.

10 An object of this invention is to provide a method for designing a new synthetic pair of connectrons in a genome that act in cooperative mode such that with respect to the individual connectrons there is an increased lifetime of connectron control of a set of genes.

15

Description of the Drawings and Tables

- Figure 1 Shows how (a) a lower connectron and (b) an upper connectron, form (c) an lower-upper connectron pair
- 5 Figure 2 Shows how (a) a left connectron and (b) a right connectron, form (c) a left-right connectron pair
- Figure 3 Shows (a) a symmetric lower-upper connectron pair, (b) an asymmetric lower-upper connectron pair
- 10 Figure 4 Shows (a) a symmetric left-right connectron pair, (b) an asymmetric left-right connectron pair
- Figure 5 Shows (a) Concise representation of an asymmetric lower-upper connectron pair and (b) detailed representation of a asymmetric lower-upper connectron pair
- 15 Figure 6 Shows (a) Concise representation of an asymmetric left-right connectron pair and (b) detailed representation of a asymmetric left-right connectron pair
- 20 Figure 7 Shows the four variations of symmetric lower-upper connectron pairs - (a) dominant - dominant, (b) anti-dominant - dominant, (c) dominant - anti-dominant, and (d) anti-dominant - dominant
- 25 Figure 8 Shows the four variations of symmetric left-right connectron pairs - (a) dominant - dominant, (b) anti-dominant - dominant, (c) dominant - anti-dominant, and (d) anti-dominant - dominant
- 30 Figure 9 Shows the four variations of asymmetric lower-upper connectron pairs - (a) dominant - dominant, (b) anti-dominant - dominant, (c) dominant - anti-dominant, and (d) anti-dominant - dominant
- 35 Figure 10 Shows the four variations of asymmetric left-right connectron pairs - (a) dominant - dominant, (b) anti-dominant - dominant, (c) dominant - anti-dominant, and (d) anti-dominant - dominant
- 40 Figure 11 Shows (a) the competitive blocking of symmetric lower-upper long-lived connectrons, (b) the competitive blocking of symmetric left-right long-lived connectrons, and (c) the relative timing windows for competitive blocking of symmetric long-lived connectrons
- 45

5 Figure 12 Shows (a) the competitive blocking of asymmetric lower-upper long-lived connectrons, (b) the competitive blocking of asymmetric lower-upper long-lived connectrons, (c) the competitive blocking of asymmetric left-right long-lived connectrons, and (d) the competitive blocking of asymmetric left-right long-lived connectrons

10 Figure 13 Shows (a) the timing windows for competitive blocking of asymmetric long-lived connectrons

15 Figure 14 Shows (a) a fully symmetric connectron tetrad, (b) non-competitive blocking effect by a left short-lived connectrons of a long-lived cooperative connectron pair, (c) non-competitive blocking effect by a right short-lived connectrons of a long-lived cooperative connectron pair, and (d) the timing windows for non-competitive blocking of asymmetric long-lived connectrons

20

Description of the Invention

The basic methods patent application for the determination of connectron structure defines the DNA and RNA sequence components that make up a connectron, as well as presenting examples of different sorts of connectrons from many different types of genomes. The computer algorithm presented in that patent application shows how to find connectrons in a particular genome. The genomic patent applications utilize the power of this computer algorithm to determine all of the connectrons in a particular genome. Although the basic methods patent application identifies permanent, transient and one-shot connectrons, the view presented is that of a single connectron. This patent application presents the relationships among pairs of connectrons. This invention will allow us to organize the connectrons in a genome and show how pairs of connectrons work together to produce new gene expression regulation properties. In particular, this invention will allow us to show how different C1/C2 connectron sequences from different gene expression events can cooperate

25

30

35

40

to form a pair of long-lived connectrons. The ability to form very specific and cooperative conjunctive events makes it possible for biological systems to form arbitrarily complex control procedures that may very well be needed for cellular differentiation and the development of a complete multi-celled organism from a single cell.

A connectron forms a loop in a piece of double-stranded DNA. As shown in figure 1a the DNA runs from the 5' end shown on the lower left in a counter-clockwise direction to the 3' end shown on the lower right. The RNA generated by the promotion and transcription of some gene or pseudo-gene somewhere in the genome binds to two distinct double-stranded DNA sequences to form two distinct triple-stranded generalized Hoogsteen helices. In figure 1a the first triple-stranded (generalized Hoogsteen) helix is called A and the second helix is called B. The A helix forms along the major groove of the DNA in the 5' to 3' direction. Similarly the B helix in figure 1a forms along the major groove of the DNA in the 5' to 3' direction. The A-B pair of triple-stranded helices occupy the lower position in the X-shape formed by the loop. Hence in figure 1a the connectron is described as a "lower connectron". In figure 1b both the A and B helices form in the 5' to 3' direction, but they occupy the upper position in the X-shape formed by the loop. Hence in figure 1b the connectron is described as an "upper connectron". In figure 1c the lower and upper connectrons are shown binding simultaneously.

In figure 2a the A helix forms along the major groove of the DNA in the 5' to 3' direction but the B helix forms along the major groove of the DNA by binding along the major groove of the double helix in the 3' to 5' direction. The importance of this connectron is that the RNA switches strands as it moves from A-helix binding to B-helix binding. This is true in all left and right connectrons. The A-B pair of helices in figure

2a occupy the left position in the X-shape formed by the loop. Hence in figure 2a the connectron is described as a "left connectron". In figure 2b the A triple-strand helix forms along the major groove of the DNA in the 3' to 5' direction, so this connectron is given the designation A-B and is described as a "right connectron". In figure 2c the left and right connectrons are shown binding simultaneously.

In figure 3a the pair of lower and upper connectrons have the same sequences (i.e. A and B) hence this pair of connectrons is called a "symmetric lower-upper connectron pair". In figure 3b the lower connectron has the sequence A-B and the upper connectron has the sequence C-D hence this pair of connectrons is called an "asymmetric lower-upper connectron pair".

In figure 4a the pair of left and right connectrons have the same sequences (i.e. A and B) hence this pair of connectrons is called a "symmetric left-right connectron pair". In figure 4b the left connectron has the sequence A-B and the right connectron has the sequence C-D, where C is not equal to A and/or D is not equal to B hence this pair of connectrons is called an "asymmetric left-right connectron pair".

Figure 5a - a re-statement of figure 3b - is a concise representation of an asymmetric lower-upper connectron pair. Figure 5b is a detailed representation of the same asymmetric lower-upper connectron pair showing the sequence relationships between the RNA strand and the two DNA strands. The equivalence of the RNA-strand sequence and the 5' to 3' DNA-strand sequence means that the RNA-strand sequence will share the hydrogen bonds to the 3' to 5' DNA-strand sequence.

Figure 6a - a re-statement of figure 4b - is a concise representation of an asymmetric left-right connectron pair.

Figure 6b is a detailed representation of the same asymmetric left-right connectron pair showing the sequence relationships between the RNA strand and the two DNA strands. The

equivalence of the RNA-strand sequence and the 5' to 3' DNA-strand sequence for the first triple helix of each of these connectrons means that the RNA-strand sequence will share the hydrogen bonds to the 3' to 5' DNA-strand sequence.

Similarly, the equivalence of the RNA-strand sequence and the 3' to 5' DNA-strand sequence for the second triple helix of each of these connectrons means that the RNA-strand sequence will share the hydrogen bonds to the 5' to 3' DNA-strand sequence.

Figure 7a shows the lower and upper connectrons both binding in the dominant direction with the sequence A-B hence this pair of connectrons is called a "dominant - dominant symmetric lower-upper connectron pair". Figure 7b shows the lower and upper connectrons both binding in the anti-dominant direction with the sequence B'-A' hence this pair of connectrons is

called an "anti-dominant - anti-dominant symmetric lower-upper connectron pair". In figure 7c the lower connectron binds in the dominant direction with the sequence A-B and the upper connectron binds in the anti-dominant direction with the sequence B'-A' hence this pair of connectrons is called a

"dominant - anti-dominant symmetric lower-upper connectron pair". In figure 7d the lower connectron binds in the anti-dominant direction with the sequence B'A' and the upper connectron binds in the dominant direction with the sequence A-B. In figure 7 each of the four sequence pairs is

different, hence there are four different types of symmetric lower-upper connectron pairs.

Figure 8a shows the left and right connectrons both binding in the dominant direction with the sequences A-B hence this pair

of connectrons is called a "dominant - dominant symmetric

left-right connectron pair". Figure 8b shows the lower and upper connectrons both binding in the anti-dominant direction with the sequence B'-A' hence this pair of connectrons is called an "anti-dominant - anti-dominant symmetric left-right connectron pair". In figure 8c the left connectron binds in the dominant direction with the sequence A-B and the right connectron binds in the anti-dominant direction with the sequence B'-A' hence this pair of connectrons is called a "dominant - anti-dominant symmetric left-right connectron pair". In figure 8d the left connectron binds in the anti-dominant direction with the sequence B'-A' and the right connectron binds in the dominant direction with the sequence A-B hence this pair of connectrons is called an "anti-dominant - dominant symmetric left-right connectron pair". In figure 8 each of the four sequence pairs is different, hence there are four different types of symmetric left-right connectron pairs.

Figure 9a shows the lower and upper connectrons both binding in the dominant direction but the sequences of the two connectrons are different. The lower connectron has the sequence A-B and the upper connectron has the sequence C-D hence this pair of connectrons is called a "dominant - dominant asymmetric lower-upper connectron pair". In figure 9b the lower connectron binds in the dominant direction with the sequence B'-A' and the upper connectron binds in the anti-dominant direction with the sequence D'-C' hence this pair of connectrons is called an "anti-dominant - anti-dominant asymmetric lower-upper connectron pair". In figure 9c the lower connectron binds in the dominant direction with the sequence A-B and the upper connectron binds in the anti-dominant direction with the sequence D'-C' hence this pair of connectrons is called a "dominant - anti-dominant asymmetric lower-upper connectron pair". In figure 9d the lower connectron binds in the anti-dominant direction with the sequence B'-A' and the upper connectron binds in the dominant

direction with the sequence C-D hence this pair of connectrons is called an "anti-dominant - dominant asymmetric lower-upper connectron pair". In figure 9 each of the four sequence pairs is different, hence there are four different types of

5 asymmetric lower-upper connectron pairs.

Figure 10a shows the left and right connectrons both binding in the dominant direction but the sequences of the two connectrons are different. The left connectron has the

10 sequence A-B and the right connectron has the sequence C-D hence this pair of connectrons is called a "dominant - dominant asymmetric left-right connectron pair". In figure 10b the left connectron binds in the anti-dominant direction with the sequence B'-A' and the right connectron binds in the

15 anti-dominant direction with the sequence D'-C' hence this pair of connectrons is called an "anti-dominant - anti-dominant asymmetric left-right connectron pair". In figure 10c the left connectron binds in the dominant direction with the sequence A-B and the right connectron binds in the anti-

20 dominant direction with the sequence D'-C' hence this pair of connectrons is called a "dominant - anti-dominant asymmetric left-right connectron. In figure 10d the left connectron binds in the anti-dominant direction with the sequence B'-A'

and the right connectron binds in the dominant direction with the sequence C-D hence this pair of connectrons is called an "anti-dominant - dominant asymmetric left-right connectron.

In figure 10 each of the four sequence pairs is different, hence there are four different types of asymmetric left-right connectron pairs.

30

The lifetime of a single connectrons is easy to understand.

Consider a single connectron as shown in figure 1a. For the sake of example let the A triple-strand (generalized

Hoogsteen) helix be the minimum length of 15 bases and let the

35 B triple-strand helix be some long length, for example 100

bases. Remember that the RNA-DNA structure of the connectrons is immersed in a bath of water at 37 degrees Celsius. Thermal motion will cause the A triple-strand helix to dissolve into the RNA and DNA components much more rapidly than the much longer B triple-strand helix, so the lifetime of the connectrons varies directly with the length of the shorter of the two triple-strand helices. If the length of the A and B helices are the same then the lifetime of the connectron varies directly with the length of either helix.

10

Now think about the lifetime of a pair of connectrons as shown in figure 1c. For the sake of simplicity, assume that the A and B helices are the same length. In order for the loop to open up, at least one lower helix and one upper helix has to dissolve at the same time. Either or both of the two lower helices can dissolve at the same time but as long as the upper pair of helices is not dissolved, the loop will stay closed. The same is true for the reverse - either of both of the two upper helices can dissolve at the same time but as long as the two lower helices stay intact, the loop stays closed.

20

Physical chemistry is replete with two-part binding events like this. The general description of such events is that the lifetime varies directly with the product of the two binding energies. In the case of a pair of lower-upper connectrons, the lifetime varies directly as the product of the shorter of the lower helices and the shorter of the upper helices. Of course, the same thing is true for the left-right connectron pairs shown in figure 2c.

25

30 Whatever the binding energy of an RNA strand is with respect to its cognate double-stranded DNA sequence, whenever a sequence pair (for example A-B or B'-A') can form a pair of connectrons, the binding energy of the pair of connectrons is the product of the binding energy of the each connectron.

35 When two different sequence pairs (for example A-B and B'-A'

as shown in figure 7) form a pair of connectrons then the pair of connectrons can be formed as one of the four following combinations

- | | | | |
|---|-----------|---------------|-------------------|
| 5 | Figure 5a | lower (A-B) | and upper (A-B) |
| | Figure 5b | lower (B'-A') | and upper (B'-A') |
| | Figure 5c | lower (A-B) | and upper (B'-A') |
| | Figure 5d | lower (B'-A') | and upper (A-B) |

10 In principle, A-B and B'-A' could be produced by the expression of two different genes. Because the pair of connectrons can form in four different ways, the two genes causing the production of the two different RNAs are competing for control of the formation of the connectron pair.

15

The left-right connectron pairs in figure 8 have the same properties as the lower-upper connectron pairs in figure 7. In principle, A-B and B'-A' could be produced by the expression of two different genes. Because the pair of connectrons can
20 form in four different ways, the two genes causing the production of the two different RNAs are competing for control of the formation of the connectron pair.

When two different sequence pairs (for example A-B and C-D)
25 form a pair of connectrons then the pair of connectrons can be formed in only one way as shown in figure 9a. A-B and C-D can be produced by the expression of two different genes. Because the pair of connectrons can form in only one way, the two genes causing the production of the two different RNAs are
30 cooperating for control of the formation of the connectron pair. The same cooperative behavior is also true of the sequence combinations in figures 9b, 9c and 9d.

Like figures 7 and 8 (that describe symmetric connectron
35 pairs), figures 9 and 10 (that describe asymmetric connectron

pairs) share the same properties. When two different sequence pairs (for example A-B and C-D) form a pair of connectrons then the pair of connectrons can be formed in only one way as shown in figure 10a. A-B and C-D can be produced by the
5 expression of two different genes. Because the pair of connectrons can form in only one way, the two genes causing the production of the two different RNAs are cooperating for control of the formation of the connectron pair. The same cooperative behavior is also true of the sequence combinations
10 in figures 10b, 10c and 10d.

In figures 7 and 8 (that describe symmetric connectron pairs) the connectron pair constructs produce competition whereas in figures 9 and 10 (that describe asymmetric connectron pairs)
15 the connectron pair constructs produce cooperation. Figures 7 and 8 are symmetric constructs whereas figures 9 and 10 are asymmetric constructs.

Whereas in figure 7 the sequence along the DNA the X-shape of
20 the crossing (i.e. either the / sequence or the \ sequence) is the sequence A-B, in figure 8 the same elements are reverse-complements.

The algorithm described in the basic methods patent
25 application finds all of the uni-polar the connectrons in a genome. This patent application describes connectrons in terms of their symmetry properties (i.e. uni-polar, bi-polar, lower, upper, left, right, symmetric, asymmetric). The original algorithm has been modified and the connectron
30 structure of the genomes recomputed to find both the uni-polar and bi-polar connectrons. The modification of the basic connectron determination algorithm to identify the left-right connectrons required only a half dozen lines of code change which is at or below the level of resolution of the flow
35 charts presented in the basic methods patent application. The

utility of this patent application is that we have shown that pairs of connectrons both compete and cooperate by forming in the same place (i.e. the X-shaped loop interaction region) to produce lifetimes that vary directly with the product of the lifetimes of the individual connectrons.

Figure 11a shows how one source (A-B) of the C1/C2 RNA that forms a lower-upper connectron pair with a relatively short product lifetime can temporally compete with another source of a much longer C1/C2 RNA which could form a much longer-lived symmetric connectron pair. Figure 11b shows how one source (A-B) of the C1/C2 RNA that forms a left-right connectron pair with a relatively short product lifetime can temporally compete with another source of a much longer C1/C2 RNA which could form a much longer-lived symmetric connectron pair. As shown in figure 11c, the shorter A-B connectron pair only has to last throughout the expression window of the longer connectron pair in order to prevent the longer-lived connectron pair from forming. After the short-lived A-B connectron pair expires, the loop is effectively open.

Figure 12a shows how a short-lived lower connectron can block the formation of a much longer-lived asymmetric connectron pair. Figure 12b shows how a short-lived upper connectron can block the formation of a much longer-lived asymmetric connectron pair. Figure 12c shows how a short-lived left connectron can block the formation of a much longer-lived asymmetric connectron pair. Figure 12d shows how a short-lived right connectron can block the formation of a much longer-lived asymmetric connectron pair.

Figure 13a shows the timing windows for the competitive blocking of an asymmetric long-lived connectron pairs as shown in figures 12a and 12c. Figure 13b shows the timing windows

for the competitive blocking of an asymmetric long-lived connectron pair pairs as shown in figures 12b and 12d.

Figure 14a shows how, in-principle, four connectrons could
5 form at a given site. Clearly not all four of these connectrons can form at this site at the same time because each connectron occupies two of the four target (T1 or T2) sites. The lower A-B and upper C-D pair can form at the same time or the left A-C' and right D'-B pair can form at the same
10 time. Figure 14b shows how a short-lived left connectron A-C' can block the formation of a much longer-lived cooperative connectron pair A-B and C-D. Figure 14c shows how a short-lived right connectron D'-B can block the formation of a much longer-lived cooperative connectron pair A-B and C-D. Figure
15 14d shows the timing chart for this type of temporal blocking. Figures 11c, 13a and 14d show three distinctly different types of temporal blocking. To someone skilled in the art it would be obvious that it does not matter whether the lower-upper or left-right connectrons are used for either the blocking or
20 blocked connectrons - as long as the relative patterns are maintained.

The utility of the connectron pairs shown in figures 1 to 14 is that they form the primitives of a language that can build
25 arbitrarily large and complex patterns of structural and temporal connectron control of gene expression. These language primitives can be used to analyze patterns of connectron control of gene expression in all types of genomes (i.e. prokaryotic, Archeal and eukaryotic). These language
30 primitives can also be used to create new patterns of connectron control of gene expression in all types of genomes. These same primitives will help us to understand how cells differentiate from each other in terms of their gene expression and how a single cell develops into a complete
35 organism.

Although figures 1 to 14 function in the first instance to describe the relationships between the control sequences (i.e. the C1s and C2s) produced by the same or different gene

5 expressions and the target sequences (i.e. the T1s and the T2s) in a pair of connectrons, these same figures can also function as the basis for the design of new synthetic pairs of connectrons. For example, the target sequences (A-B) that form the symmetric connectron pair shown in figure 3a could be
10 modified by changing the upper connectron T1 sequence from A to C and the upper connectron T2 sequence from B to D to form the asymmetric connectron pair shown in figure 3b. The C1/C2 sequences C-D could then be inserted in the 3'UTR of some gene so the A-B and C-D connectron pair would be formed only when
15 two genes expressed. This modification of the upper connectron sequences is an example of how all the connectron pair properties in figures 1 to 14 could be instantiated either by de-novo sequence placement or by partial modification of existing sequences and relationships. Anyone
20 skilled in the art should be able to convert the descriptions of connectron-pair properties in figures 1 to 14 into design specifications thereby opening up the control of gene expression to a whole range of new possibilities.

25 The utility of pairs of connectrons (a) whether existing or designed or (b) whether competitive or cooperative is that the lifetime of a single connectron whether it is short or long is multiplied by the existence of an adjacent connectron of similar or different lifetime properties. While the product
30 of the lifetimes of two 15-base connectrons is a modest 225, the product of the lifetimes of two 100-base connectrons would provide an impressive 10,000. Long-lived connectron pairs

provide the possibility of turning off a set of genes for extended periods of time. In the examples that follow, Nature has used sequence matches that vary in this range.

Examples

Figures 1 to 14 provide a large number of ways of describing and designing connectron pairs in a genome. We give examples of the description of symmetric and asymmetric connectron pairs in six classes of genomes (prokaryotic, Archeal, single-celled eukaryotic, multi-celled eukaryotic, mammalian and plant). We also give two examples of the design of an asymmetric connectron pair in a single-celled eukaryote and a mammal. It is clear that many other variations of symmetric and asymmetric connectron pairs could be described or designed by someone skilled in the art.

Description of a symmetric lower-upper connectron pair in *E. coli*.

E. coli is a prokaryotic organism. A single connectron has been selected from the *E. coli* connectrome to illustrate the properties of a lower-upper connectron pair. Because the connectron is very long it can be split into two connectrons that then bind as a pair. In this and each of the following examples, a header indicates the function of each data field. Because of print-page limitations, the "sequence of match" field has been moved to the left side of each example.

The connectron 1434 has a C1-T1 binding length of 182 bases and a C2-T2 binding length of 171 bases. The shorter of the two matches of 171 bases is then halved with the first half becoming the A and the second half becoming the B in figure 3a producing a producted connectron pair lifetime of 7225.

genome
Connectron id
chromosome
contig
(.groups) id
type
CP = control element on positive strand
CN = control element on negative strand
TP = target element on positive strand
TN = target element on negative strand
match start
match stop
type of Connectron
l/u = lower/upper
l/r = left/right
source of Connectron
g = gene
p = pseudogene
length of match
sequence of match
eco 1434 1 1 7435 CP 4505.026 4505.207 l/u g 182

CTGTAGATTCAATCTGTCAATGCAACACCCCTTTCAATTATCTCTTTCGG
TGTTTTGAACTTCAGTGTCTTTCTCGGTCTGTTGTTAGCTGAGCAGCAA
CCAGATCTAGTTCATGTTGAGTATATTGGGCAAGACATGTCTTTTAGGA
AAGTACTGCCGAATTAGCCCATTTGTGTTCTC

eco 1434 1 1 508 TN 279.155 279.336 l/u g 182

CTGTAGATTCAATCTGTCAATGCAACACCCCTTTCAATTATCTCTTTCGG
TGTTTTGAACTTCAGTGTCTTTCTCGGTCTGTTGTTAGCTGAGCAGCAA
CCAGATCTAGTTCATGTTGAGTATATTGGGCAAGACATGTCTTTTAGGA
AAGTACTGCCGAATTAGCCCATTTGTGTTCTC

eco 1434 1 1 7435 CP 4505.031 4505.201 l/u g 171

GATTCAATCTGTCAATGCAACACCCCTTTCAATTATCTCTTTCGGTGT
TGAACCTTCAGTGTCTTTCTCGGTCTGTTGTTAGCTGAGCAGCAACCAGA
TCTAGTTCATGTTGAGTATATTGGGCAAGACATGTCTTTTAGGAAAGTA
CTGCCGAATTAGCCCATTTGT

eco 1434 1 1 472 TN 270.811 270.981 l/u g 171

GATTCAATCTGTCAATGCAACACCCCTTTCAATTATCTCTTTCGGTGT
TGAACCTTCAGTGTCTTTCTCGGTCTGTTGTTAGCTGAGCAGCAACCAGA
TCTAGTTCATGTTGAGTATATTGGGCAAGACATGTCTTTTAGGAAAGTA
CTGCCGAATTAGCCCATTTGT

Can form an AB symmetric pair of l/u Connectrons with a lifetime = 85 X 85 = 7225

171

GATTCAATCTGTCAATGCAACACCCCTTTCAATTATCTCTTTCGGTGT
TGAACCTTCAGTGTCTTTCTCGGTCTGTTGTTAGCTGAGCAGCAACCAGA
TCTAGTTCATGTTGAGTATATTGGGCAAGACATGTCTTTTAGGAAAGTA
CTGCCGAATTAGCCCATTTGT

171

GATTCAATCTGTCAATGCAACACCCCTTTCAATTATCTCTTTCGGTGT
TGAACCTTCAGTGTCTTTCTCGGTCTGTTGTTAGCTGAGCAGCAACCAGA
TCTAGTTCATGTTGAGTATATTGGGCAAGACATGTCTTTTAGGAAAGTA
CTGCCGAATTAGCCCATTTGT

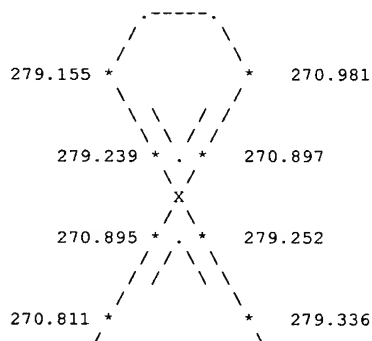
279.155 279.239 279.252 279.336 --- 270.811 270.895 270.897 270.981

65

70

75

80



**Description of a symmetric lower-upper connectron pair
in *S. tokodaii***

S. tokodaii is a Archeal organism. In this and the following
5 examples, the header does not show all the cases for a given
data field.

The connectron 4240 has a C1-T1 binding length of 67 bases and a
C2-T2 binding length of 85 bases. The effective match of 52
10 bases is then halved with the first half becoming the A and the
second half becoming the B in figure 3a producing a producted
connectron pair lifetime of 676.

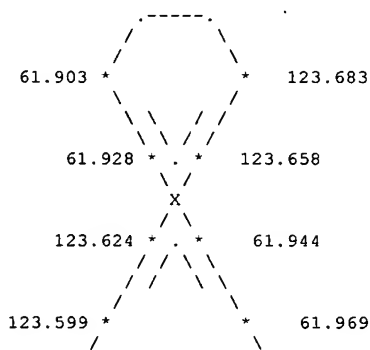
genome	Connectron id	chromosome	contig	(.groups) id	type	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match										
sto	4240	1	1	3986	CN	1178.996	1179.062	1/u	g	67
TGTACCCCTTCAAGTAAGCCTCATTTAAGGGAGTTTCTCCCTTGAATA AACTACCGGTACATGA										
sto	4240	1	1	447	TP	61.903	61.969	1/u	g	67
TGTACCCCTTCAAGTAAGCCTCATTTAAGGGAGTTTCTCCCTTGAATA AACTACCGGTACATGA										
sto	4240	1	1	3986	CN	1178.963	1179.047	1/u	g	85
TTGTAATATTATATCAGTTTACTTCTAATATACTGTACCCCTTCAAGTA AGCCTCATTTAAGGGAGTTTCTCCCTTGAATAAA										
sto	4240	1	1	646	TP	123.599	123.683	1/u	g	85
TTGTAATATTATATCAGTTTACTTCTAATATACTGTACCCCTTCAAGTA AGCCTCATTTAAGGGAGTTTCTCCCTTGAATAAA										

Can form an AB symmetric pair of 1/u Connectrons with a lifetime = 26 X 26 = 676

TGTACCCCTTCAAGTAAGCCTCATTTAAGGGAGTTTCTCCCTTGAATA
AA

TGTACCCCTTCAAGTAAGCCTCATTTAAGGGAGTTTCTCCCTTGAATA
AA

61.903 61.928 61.944 61.969 --- 123.599 123.624 123.658
123.683



Description of a symmetric lower-upper connectron pair in *S. cerevisiae*

The connectron 385 has a C1-T1 binding length of 117 and a C2-T2 binding length also of 117 bases. Since the two matches are equal, the 117 bases are then halved with the first half becoming the A and the second half becoming the B in figure 3a producing a producted connectron pair lifetime of 3364.

S. cerevisiae is a single-celled eukaryotic organism.

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match									
yst	385	15	15	28455 CP	975.950	976.066	1/u g		117
TTACTAGTATATTATCATATACGGTGTAGAAAGATGACGCAAATGATGAG AAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGATAATGTA ATAGGATCAATGAATAT									
yst	385	1	1	419 TN	165.888	166.004	1/u g		117
TTACTAGTATATTATCATATACGGTGTAGAAAGATGACGCAAATGATGAG AAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGATAATGTA ATAGGATCAATGAATAT									
yst	385	15	15	28455 CP	975.950	976.066	1/u g		117
TTACTAGTATATTATCATATACGGTGTAGAAAGATGACGCAAATGATGAG									

AAATAGTCATCTAAATTAGTGGAGCTGAAACGCAAGGATTGATAATGTA
 ATAGGATCAATGAATAT

5 yst 385 1 1 355 TN 160.257 160.373 1/u g 117

TTACTAGTATATTATCATATACGGTGTAGAGATGACGCAAATGATGAG
 AAATAGTCATCTAAATTAGTGGAGCTGAAACGCAAGGATTGATAATGTA
 ATAGGATCAATGAATAT

10 Can form an AB symmetric pair of 1/u Connectrons with a lifetime = 58 X 58 = 3364

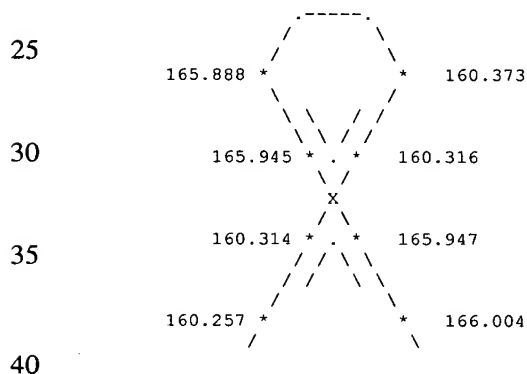
TTACTAGTATATTATCATATACGGTGTAGAGATGACGCAAATGATGAG
 AAATAGTCATCTAAATTAGTGGAGCTGAAACGCAAGGATTGATAATGTA
 ATAGGATCAATGAATAT

15 117

TTACTAGTATATTATCATATACGGTGTAGAGATGACGCAAATGATGAG
 AAATAGTCATCTAAATTAGTGGAGCTGAAACGCAAGGATTGATAATGTA
 ATAGGATCAATGAATAT

20 117

165.888 165.945 165.947 166.004 --- 160.257 160.314 160.316 160.373

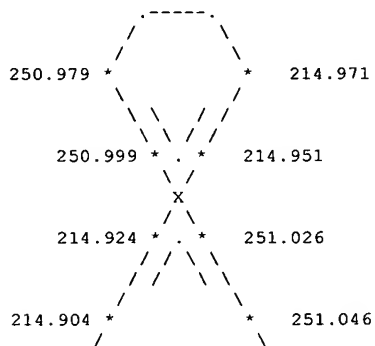


Description of a symmetric lower-upper connectron pair in *C. elegans*

45 *C. elegans* is a 1,000-celled eukaryotic organism.

The connectron 55 has a C1-T1 binding length of 68 and a C2-T2
 binding length also of 68 bases. The effective match of 43
 50 bases is then halved with the first half becoming the A and the
 second half becoming the B in figure 3a producing a producted
 connectron pair lifetime of 441.

genome	Connectron id	chromosome	contig	(.groups) id	type	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match										
wrn	55	1	1	380	CN	221.205	221.272	1/u g		68
GGGAATTGCTTCGTCAAATGATCGACGGAGGGCTTTTGGCCATCTGCAAG										
GATAAACTCGCATGTCGA										
wrn	55	1	1	433	TN	250.979	251.046	1/u g		68
GGGAATTGCTTCGTCAAATGATCGACGGAGGGCTTTTGGCCATCTGCAAG										
GATAAACTCGCATGTCGA										
wrn	55	1	1	380	CN	221.180	221.247	1/u g		68
GAGCTCGCAACACCGGCCGAGCAGCGGGAATTGCTTCGTCAAATGATCGA										
CGGAGGGCTTTTGGCCAT										
wrn	55	1	1	354	TN	214.904	214.971	1/u g		68
GAGCTCGCAACACCGGCCGAGCAGCGGGAATTGCTTCGTCAAATGATCGA										
CGGAGGGCTTTTGGCCAT										
Can form an AB symmetric pair of 1/u Connectrons with a lifetime = 21 X 21 = 441										
GGGAATTGCTTCGTCAAATGATCGACGGAGGGCTTTTGGCCAT										43
GGGAATTGCTTCGTCAAATGATCGACGGAGGGCTTTTGGCCAT										43
250.979 250.999 251.026 251.046 --- 214.904 214.924 214.951 214.971										



Description of a symmetric lower-upper connectron pair in *H. sapiens*

H. sapiens is a multi-celled eukaryotic organism - a mammal.

The connectron 1211 has a C1-T1 binding length of 58 bases and a C2-T2 binding length also of 58 bases. Since the two matches are equal, 58 is then halved with the first half becoming the A and the second half becoming the B in figure 3a producing a producted connectron pair lifetime of 841.

genome	Connectron id	chromosome	contig	(.groups) id	type	match start	match stop	type of Connectron	source of Connectron	length of match
hsd	1211	4	1	1331	CP	16.381	16.438	l/u	g	58

GGTGAGTACCTTTCTATGAAGGTGATAAGGATCCACTGAGTCTTCCATAT
AAAGATCA

hsd	1211	4	1	1542	TP	500.217	500.274	l/u	g	58
-----	------	---	---	------	----	---------	---------	-----	---	----

GGTGAGTACCTTTCTATGAAGGTGATAAGGATCCACTGAGTCTTCCATAT
AAAGATCA

hsd	1211	4	1	1331	CP	16.381	16.438	l/u	g	58
-----	------	---	---	------	----	--------	--------	-----	---	----

GGTGAGTACCTTTCTATGAAGGTGATAAGGATCCACTGAGTCTTCCATAT
AAAGATCA

hsd	1211	4	1	1559	TP	504.937	504.994	l/u	g	58
-----	------	---	---	------	----	---------	---------	-----	---	----

GGTGAGTACCTTTCTATGAAGGTGATAAGGATCCACTGAGTCTTCCATAT
AAAGATCA

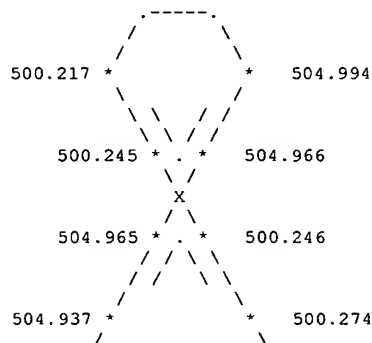
Can form an AB symmetric pair of l/u Connectrons with a lifetime = 29 X 29 = 841

GGTGAGTACCTTTCTATGAAGGTGATAAGGATCCACTGAGTCTTCCATAT
AAAGATCA

GGTGAGTACCTTTCTATGAAGGTGATAAGGATCCACTGAGTCTTCCATAT AAAGATCA	58
--	----

GGTGAGTACCTTTCTATGAAGGTGATAAGGATCCACTGAGTCTTCCATAT
AAAGATCA

500.217 500.245 500.246 500.274 --- 504.937 504.965 504.966 504.994



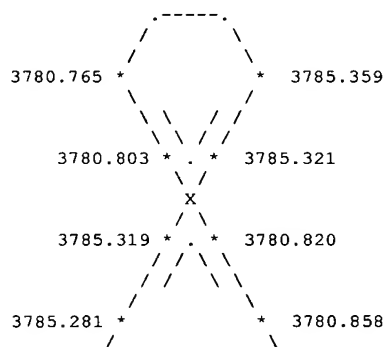
Description of a symmetric lower-upper connectron pair in *A. thaliana*

A. thaliana is a multi-celled eukaryotic organism - a plant.

The connectron 3 has a C1-T1 binding length of 94 bases and a C2-T2 binding length of 79 bases. The shorter of the two matches of 79 bases is then halved with the first half becoming

the A and the second half becoming the B in figure 5a producing a produced connectron pair lifetime of 1521.

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match				type					
ath	3	5	1	29822 CN	21590.870	21590.960	1/u g		94
20 TGTGAAAGTTAAACTTGATTTTGAATCAAGTTTAATTATTGGATCAATT ATCCAATAATTAATTATGGCCAAATCCAAGTTCTAGAGTTTCT									
ath	3	1	1	7951 TP	3780.765	3780.858	1/u g		94
25 TGTGAAAGTTAAACTTGATTTTGAATCAAGTTTAATTATTGGATCAATT ATCCAATAATTAATTATGGCCAAATCCAAGTTCTAGAGTTTCT									
ath	3	5	1	29822 CN	21590.870	21590.950	1/u g		79
30 TGTGAAAGTTAAACTTGATTTTGAATCAAGTTTAATTATTGGATCAATT ATCCAATAATTAATTATGGCCAAATCCAA									
ath	3	1	1	7985 TP	3785.281	3785.359	1/u g		79
35 TGTGAAAGTTAAACTTGATTTTGAATCAAGTTTAATTATTGGATCAATT ATCCAATAATTAATTATGGCCAAATCCAA									
Can form an AB symmetric pair of 1/u Connectrons with a lifetime = 39 X 39 = 1521									
40 TGTGAAAGTTAAACTTGATTTTGAATCAAGTTTAATTATTGGATCAATT ATCCAATAATTAATTATGGCCAAATCCAA									
45 TGTGAAAGTTAAACTTGATTTTGAATCAAGTTTAATTATTGGATCAATT ATCCAATAATTAATTATGGCCAAATCCAA									
3780.765	3780.803	3780.820	3780.858	---	3785.281	3785.319	3785.321		
3785.359									



Description of an asymmetric lower-upper connectron pair in *E. coli*

The connectron 14918 has a C1-T1 binding length of 27 bases and a C2-T2 binding length of 35 bases. The shorter of the two matches at 27 bases produces the lifetime for this connectron. The connectron 15118 has a C1-T1 binding length of 20 bases and

a C2-T2 binding length of 22 bases. The shorter of the two matches at 20 bases produces the lifetime for this connectron. The lifetime of this pair of anti-dominant - dominant connectrons as shown in figure 9d is 540.

5

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match				type					
eco 14918	1	1	7316 CN	4454.807	4454.833	l/u g			27
AAATGCCGGATGCGGCGTGAACGCCTT									
eco 14918	1	1	6955 TN	4242.757	4242.783	l/u g			27
AAATGCCGGATGCGGCGTGAACGCCTT									
eco 14918	1	1	7316 CN	4454.810	4454.844	l/u g			35
TGCCGGATGCGGCGTGAACGCCTTATCCGGCCTAC									
eco 14918	1	1	6937 TN	4233.017	4233.051	l/u g			35
TGCCGGATGCGGCGTGAACGCCTTATCCGGCCTAC									

eco 15118	1	1	1544 CP	831.575	831.594	l/u g			20
TGTAGGCCGGATAAGGCGTT									
eco 15118	1	1	6939 TP	4232.999	4233.018	l/u g			20
TGTAGGCCGGATAAGGCGTT									
eco 15118	1	1	1544 CP	831.596	831.617	l/u g			22
ACGCCGCATCCGGCATTTCACA									
eco 15118	1	1	6957 TP	4242.783	4242.804	l/u g			22
ACGCCGCATCCGGCATTTCACA									

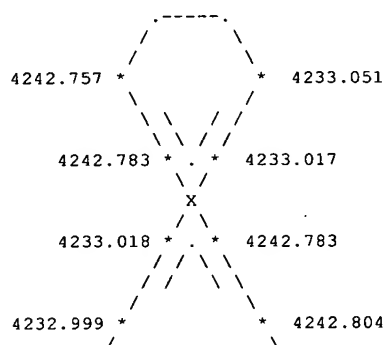
Found L/U AD AB - CD Connectron pair for 14918 and 15118 with a lifetime = 27 X 20 = 540									
.000 .001									
4242.757 4242.783 4233.017 4233.051 --- 4232.999 4233.018 4242.783 4242.804									

60

65

70

75



**Description of an asymmetric lower-upper connectron pair
in *S. tokodaii***

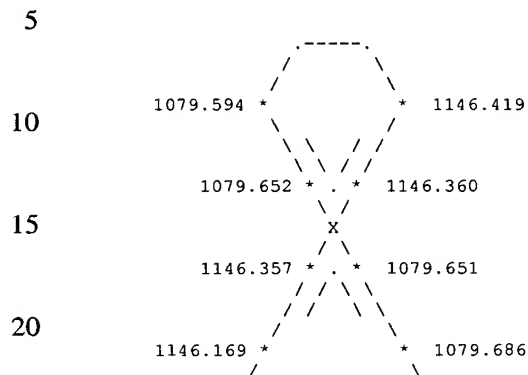
The connectron 6416 has a C1-T1 binding length of 59 bases and a
5 C2-T2 binding length of 60 bases. The shorter of the two
matches at 59 bases produces the lifetime for this connectron.
The connectron 6477 has a C1-T1 binding length of 189 bases and
a C2-T2 binding length of 36 bases. The shorter of the two
10 matches at 36 bases produces the lifetime for this connectron.
The lifetime of this pair of dominant - anti-dominant
connectrons as shown in figure 9d is 2124.

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match				type					
sto 6416	1	1	3245 CP	1036.523	1036.581	1/u g			59
ACTCCAGTGAGGGATAGGGGTAACGGACTGAAGACCCAGCCCGTGGTCT ACCGCTGGA									
sto 6416	1	1	3439 TP	1079.594	1079.652	1/u g			59
ACTCCAGTGAGGGATAGGGGTAACGGACTGAAGACCCAGCCCGTGGTCT ACCGCTGGA									
sto 6416	1	1	3250 CP	1036.635	1036.694	1/u g			60
ATGAAGGTGGTAAACCACAAACCTATGAACCGCCCTAAGGGAACCCCTCGC CCTTTAGGGC									
sto 6416	1	1	3714 TP	1146.360	1146.419	1/u g			60
ATGAAGGTGGTAAACCACAAACCTATGAACCGCCCTAAGGGAACCCCTCGC CCTTTAGGGC									

sto 6477	1	1	618 CN	120.361	120.549	1/u g			189
CTATCCCTCACCAAGAGTTGCCCTCTGCTCTTGGCTCTTGGGGACTCGGG GATATGTAGTTCTGTGCGGGGACACATATCTTCAGTATGCCACCTTTGT GGGCTTCCCCGCACTTTATTAATAGTTTAAAGCTAAGATTAAAACTTTA CCCCCGCTTAAAAGGCGAGGCTTGCCCCGCGTTTGTCA									
sto 6477	1	1	3709 TN	1146.169	1146.357	1/u g			189
CTATCCCTCACCAAGAGTTGCCCTCTGCTCTTGGCTCTTGGGGACTCGGG GATATGTAGTTCTGTGCGGGGACACATATCTTCAGTATGCCACCTTTGT GGGCTTCCCCGCACTTTATTAATAGTTTAAAGCTAAGATTAAAACTTTA CCCCCGCTTAAAAGGCGAGGCTTGCCCCGCGTTTGTCA									
sto 6477	1	1	622 CN	120.590	120.625	1/u g			36
CACCCACCCCGCTCCGTTTCGTCCAGCGGTAGACCAC									
sto 6477	1	1	3446 TN	1079.651	1079.686	1/u g			36
CACCCACCCCGCTCCGTTTCGTCCAGCGGTAGACCAC									

Found L/U DA AB - CD Connectron pair for 6416 and 6477 with a lifetime = 59 X 36 = 2124									

.001 .003
1079.594 1079.652 1146.360 1146.419 --- 1146.169 1146.357 1079.651 1079.686



Description of an asymmetric lower-upper connectron pair in *S. cerevisiae*

The connectron 3814 has a C1-T1 binding length of 72 bases and a C2-T2 binding length of 72 bases. The either of the two matches at 72 bases produces the lifetime for this connectron. The connectron 3847 has a C1-T1 binding length of 81 bases and a C2-T2 binding length of 89 bases. The shorter of the two matches at 81 bases produces the lifetime for this connectron. The lifetime of this pair of anti-dominant - dominant connectrons as shown in figure 9d is 5832.

genome	Connectron id	chromosome	contig	(.groups) id	type	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match										
yst	3814	13	13	23498	CP	362.701	362.772	1/u g		72
ATGGAATCTATATTTCTACATACTAATATTACGATTATTCCTCATTCGGT										
TTTATATGTTTCATTATCCTAT										
yst	3814	2	2	1896	TN	265.267	265.338	1/u g		72
ATGGAATCTATATTTCTACATACTAATATTACGATTATTCCTCATTCGGT										
TTTATATGTTTCATTATCCTAT										
yst	3814	13	13	23498	CP	362.701	362.772	1/u g		72
ATGGAATCTATATTTCTACATACTAATATTACGATTATTCCTCATTCGGT										
TTTATATGTTTCATTATCCTAT										
yst	3814	2	2	1495	TN	226.820	226.891	1/u g		72
ATGGAATCTATATTTCTACATACTAATATTACGATTATTCCTCATTCGGT										
TTTATATGTTTCATTATCCTAT										

```

5 yst      3847      13      13      23551 CN      372.772      372.852 1/u g      81
  AACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAAT
  ATAGATTCCATTTTGAGGATTCCTATATCCT

10 yst      3847      2      2      1496 TP      226.739      226.819 1/u g      81
  AACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAAT
  ATAGATTCCATTTTGAGGATTCCTATATCCT

15 yst      3847      13      13      23551 CN      372.836      372.924 1/u g      89
  GAGGATTCCTATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTA
  ATATTATAGCCTTTATCAACAATGGAATCCCAACAATTA

20 yst      3847      2      2      1923 TP      265.340      265.428 1/u g      89
  GAGGATTCCTATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTA
  ATATTATAGCCTTTATCAACAATGGAATCCCAACAATTA

```

```

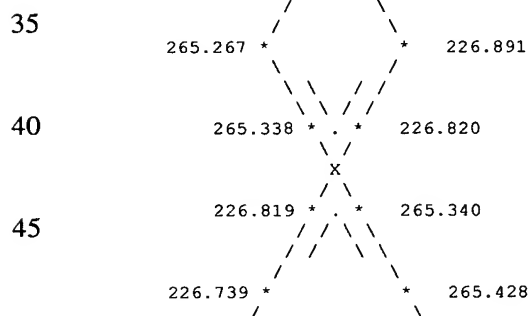
25 -----
  Found L/U AD AB - CD Connectron pair for 3814 and 3847 with a lifetime = 72 X 81 = 5832
      .002      .001

```

```

30 265.267 265.338 226.820 226.891 --- 226.739 226.819 265.340 265.428

```



Description of an asymmetric lower-upper connectron pair in *C. elegans*

```

55 The connectron 23175 has a C1-T1 binding length of 15 bases and
    a C2-T2 binding length of 18 bases. The shorter of the two
    matches at 15 bases produces the lifetime for this connectron.
    The connectron 23179 has a C1-T1 binding length of 16 bases and
    a C2-T2 binding length of 19 bases. The shorter of the two
60 matches at 16 bases produces the lifetime for this connectron.
    The lifetime of this pair of dominant - anti-dominant
    connectrons as shown in figure 9c is 240.

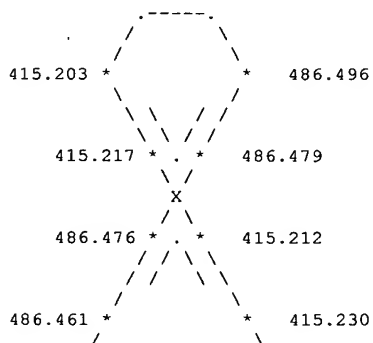
```


	genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
	sequence of match				type					
5	wrn 23175	4	2	22854	CP	708.778	708.792	l/u g	15	
10	TGGTCTGCTAAATCG									
15	wrn 23175	4	2	21925	TP	415.203	415.217	l/u g	15	
20	TGGTCTGCTAAATCG									
25	wrn 23175	4	2	22854	CP	708.793	708.810	l/u g	18	
30	AAACTTGTAGTTTGTAGT									
35	wrn 23175	4	2	22166	TP	486.479	486.496	l/u g	18	
40	AAACTTGTAGTTTGTAGT									
45	-----									
50	wrn 23179	4	2	24763	CN	1373.569	1373.584	l/u g	16	
55	ATTTAGCAGACCCAAA									
60	wrn 23179	4	2	22165	TN	486.461	486.476	l/u g	16	
65	ATTTAGCAGACCCAAA									
70	wrn 23179	4	2	24763	CN	1373.554	1373.572	l/u g	19	
75	AAACTACAAATTTTCGATTT									
80	wrn 23179	4	2	21926	TN	415.212	415.230	l/u g	19	
85	AAACTACAAATTTTCGATTT									
90	-----									

Found L/U DA AB - CD Connectron pair for 23175 and 23179 with a lifetime = 15 X 16 = 240

.005 .003

415.203 415.217 486.479 486.496 --- 486.461 486.476 415.212 415.230



Description of an asymmetric lower-upper connectron pair in *H. sapiens*

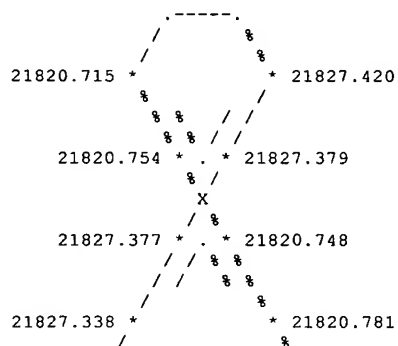
The connectron 383992 has a C1-T1 binding length of 39 bases and
a C2-T2 binding length of 41 bases. The shorter of the two
matches at 39 bases produces the lifetime for this connectron.
The connectron 383993 has a C1-T1 binding length of 40 bases and
a C2-T2 binding length of 34 bases. The shorter of the two
matches at 34 bases produces the lifetime for this connectron.
The lifetime of this pair of dominant - anti-dominant
connectrons as shown in figure 9c is 1326.

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match				type					
hsd 383992	920	19	756303	CP	21789.055	21789.092	u/d g		39
AGCCCGAGCCCCACCTCTCCCTTAGGGACCTCCGCCCAC									
hsd 383992	920	19	756563	TP	21820.715	21820.754	u/d g		39
AGCCCGAGCCCCACCTCTCCCTTAGGGACCTCCGCCCAC									
hsd 383992	920	19	756303	CP	21789.080	21789.121	u/d g		41
ACCTCCGCCCCACCTACCTCAAGCCAGGATGCCCGGAGCG									
hsd 383992	920	19	756615	TP	21827.379	21827.420	u/d g		41
ACCTCCGCCCCACCTACCTCAAGCCAGGATGCCCGGAGCG									

hsd 383993	920	19	756433	CN	21808.781	21808.820	u/d g		40
CCTAAGGGAGAGGTGGGGCTCGGGCTGAATCCCTCGTTGG									
hsd 383993	920	19	756614	TN	21827.338	21827.377	u/d g		40
CCTAAGGGAGAGGTGGGGCTCGGGCTGAATCCCTCGTTGG									
hsd 383993	920	19	756433	CN	21808.740	21808.773	u/d g		34
GCTCCGGGCATCCTGGCTTGAGGGTAGAGTGGGC									
hsd 383993	920	19	756564	TN	21820.748	21820.781	u/d g		34
GCTCCGGGCATCCTGGCTTGAGGGTAGAGTGGGC									

Found L/U DA AB - CD Connectron pair for 383992 and 383993 with a lifetime = 39 X 34 = 1326
0.006 0.002

21820.715 21820.754 21827.379 21827.420 --- 21827.338 21827.377 21820.748 21820.781



Description of an asymmetric lower-upper connectron pair in *A. thaliana*

The connectron 188312 has a C1-T1 binding length of 20 bases and a C2-T2 binding length of 30 bases. The shorter of the two matches at 20 bases produces the lifetime for this connectron. The connectron 188397 has a C1-T1 binding length of 30 bases and a C2-T2 binding length of 16 bases. The shorter of the two matches at 16 bases produces the lifetime for this connectron. The lifetime of this pair of dominant - anti-dominant connectrons as shown in figure 9c is 340.

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match									
ath 188312	18	4	269631	CP	5320.517	5320.536	u/d g		20
TTGTAGACGTATGGTGGTGG									
ath 188312	18	4	269507	TP	5311.160	5311.179	u/d g		20
TTGTAGACGTATGGTGGTGG									
ath 188312	18	4	269631	CP	5320.519	5320.548	u/d g		30
GTAGACGTATGGTGGTGGGAGACTTGTA									
ath 188312	18	4	269890	TP	5340.361	5340.390	u/d g		30
GTAGACGTATGGTGGTGGGAGACTTGTA									
ath 188397	18	4	269741	CN	5324.883	5324.921	u/d g		39

GCTCTCCACCACCACCATACTACAGTCCATCTCCAAAGG

ath 188397 18 4 269881 TN 5340.322 5340.360 u/d g 39

GCTCTCCACCACCACCATACTACAGTCCATCTCCAAAGG

ath 188397 18 4 269741 CN 5324.867 5324.882 u/d g 16

CCACCATACGTCTACA

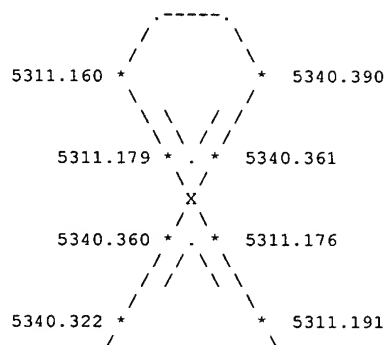
ath 188397 18 4 269509 TN 5311.176 5311.191 u/d g 16

CCACCATACGTCTACA

Found L/U DA AB - CD Connectron pair for 188312 and 188397 with a lifetime = 20 X 16 = 320

0.003 0.001

5311.160 5311.179 5340.361 5340.390 --- 5340.322 5340.360 5311.176 5311.191



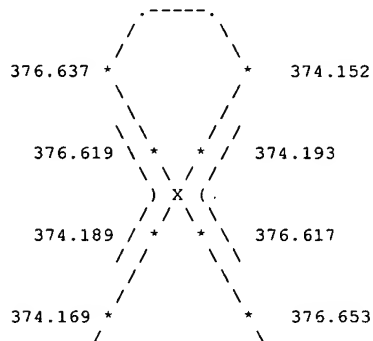
Description of an asymmetric left-right connectron pair in E. coli

The connectron 3707 has a C1-T1 binding length of 21 bases and a C2-T2 binding length of 19 bases. The shorter of the two matches at 19 bases produces the lifetime for this connectron.

The connectron 3763 has a C1-T1 binding length of 42 bases and a C2-T2 binding length of 37 bases. The shorter of the two matches at 37 bases produces the lifetime for this connectron.

The lifetime of this pair of dominant - dominant connectrons as shown in figure 10a is 703.

	genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connection	length of match
	sequence of match				type			source of Connection	
5	eco	3707	1	1	3906 CN	2338.350	2338.370	l/r g	21
	AACGCCTTATCCGGCCTACGG								
10	eco	3707	1	1	689 TP	374.169	374.189	l/r g	21
	AACGCCTTATCCGGCCTACGG								
15	eco	3707	1	1	3906 CN	2338.380	2338.398	l/r g	19
	GTAGGCCTGATAAGACGCG								
20	eco	3707	1	1	707 TN	376.619	376.637	l/r g	19
	GTAGGCCTGATAAGACGCG								
25	-----								
30	eco	3763	1	1	709 CP	376.712	376.753	l/r g	42
	GTAGGCCGGATAAGGCGTTCACGCCGCATCCGGCAGTCGTGC								
35	eco	3763	1	1	690 TN	374.152	374.193	l/r g	42
	GTAGGCCGGATAAGGCGTTCACGCCGCATCCGGCAGTCGTGC								
40	eco	3763	1	1	709 CP	376.717	376.753	l/r g	37
	CCGGATAAGGCGTTCACGCCGCATCCGGCAGTCGTGC								
45	eco	3763	1	1	706 TP	376.617	376.653	l/r g	37
	CCGGATAAGGCGTTCACGCCGCATCCGGCAGTCGTGC								
50	-----								
	Found L/R DD AB - CD Connectron pair for 3707 and 3763 with a lifetime = 19 X 37 = 703								
	.004	.002							
	374.169	374.189	376.619	376.637	---	374.152	374.193	376.617	376.653



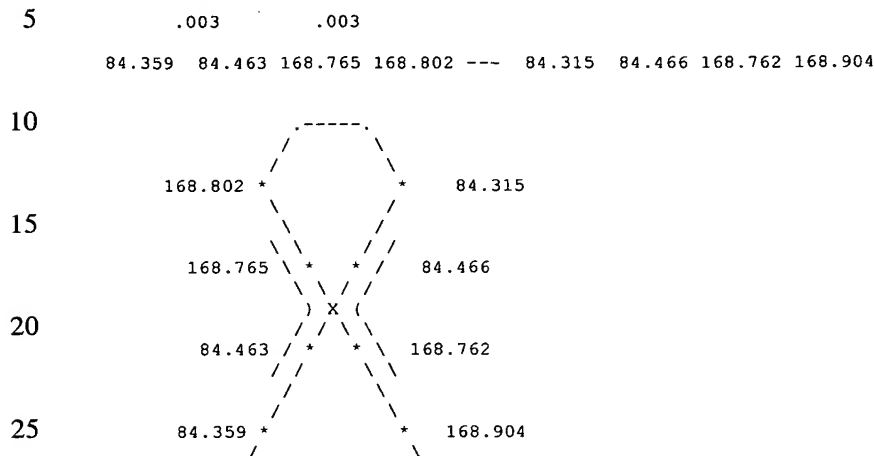
**Description of an asymmetric left-right connectron pair
in *S. cerevisiae***

The connectron 6834 has a C1-T1 binding length of 105 bases and
5 a C2-T2 binding length of 38 bases. The shorter of the two
matches at 38 bases produces the lifetime for this connectron.
The connectron 6944 has a C1-T1 binding length of 152 bases and
a C2-T2 binding length of 143 bases. The shorter of the two
matches at 143 bases produces the lifetime for this connectron.
10 The lifetime of this pair of dominant - anti-dominant
connectrons as shown in figure 10c is 5434.

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match				type					
yst 6834	7	7	10928 CN	111.321	111.425	1/r g			105
CGGTGTTAGAAGATGACGCAAATGATGAGAAATAGTCATCTAAATTAGTG									
GAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATTAAAC									
ATATA									
yst 6834	3	3	2988 TP	84.359	84.463	1/r g			105
CGGTGTTAGAAGATGACGCAAATGATGAGAAATAGTCATCTAAATTAGTG									
GAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATTAAAC									
ATATA									
yst 6834	7	7	10945 CN	111.449	111.486	1/r g			38
TCATCTACTAACTAGTATTTACGTTACTAGTATATTAT									
yst 6834	3	3	3500 TN	168.765	168.802	1/r g			38
TCATCTACTAACTAGTATTTACGTTACTAGTATATTAT									

yst 6944	4	4	5116 CN	645.641	645.792	1/r g			152
TCATCTACTAACTAGTATTTACGTTACTAGTATATTATCATATACGGTGT									
TAGAAGATGACGCAAATGATGAGAAATAGTCATCTAAATTAGTGGAAGCT									
GAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATTAAACATATAA									
AA									
yst 6944	3	3	2991 TP	84.315	84.466	1/r g			152
TCATCTACTAACTAGTATTTACGTTACTAGTATATTATCATATACGGTGT									
TAGAAGATGACGCAAATGATGAGAAATAGTCATCTAAATTAGTGGAAGCT									
GAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATTAAACATATAA									
AA									
yst 6944	4	4	5116 CN	645.641	645.783	1/r g			143
TCATCTACTAACTAGTATTTACGTTACTAGTATATTATCATATACGGTGT									
TAGAAGATGACGCAAATGATGAGAAATAGTCATCTAAATTAGTGGAAGCT									
GAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATTAA									
yst 6944	3	3	3496 TN	168.762	168.904	1/r g			143
TCATCTACTAACTAGTATTTACGTTACTAGTATATTATCATATACGGTGT									
TAGAAGATGACGCAAATGATGAGAAATAGTCATCTAAATTAGTGGAAGCT									
GAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATTAA									

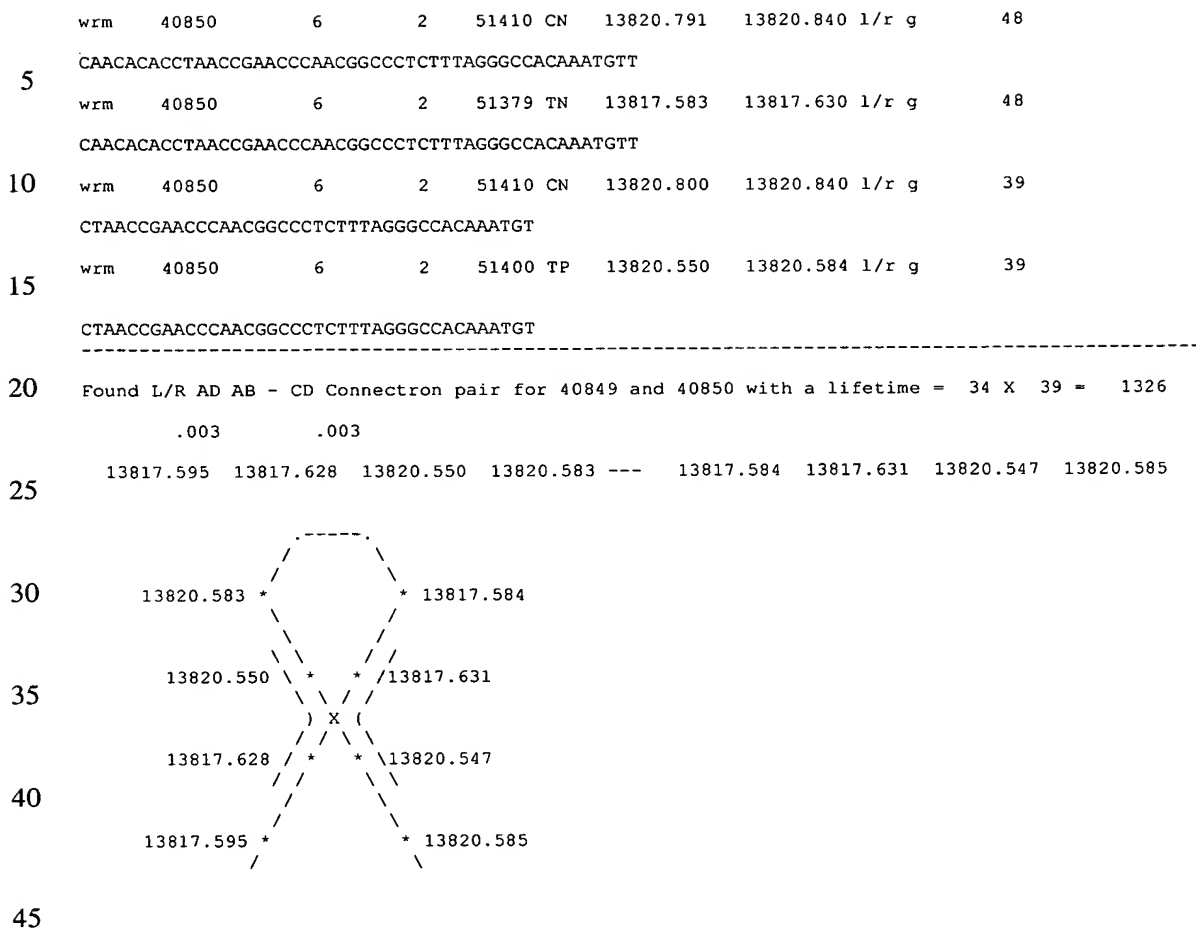
Found L/R DA AB - CD Connectron pair for 6834 and 6944 with a lifetime = 38 X 143 = 5434



30 Description of an asymmetric left-right connectron pair in *C. elegans*

The connectron 40849 has a C1-T1 binding length of 34 bases and
a C2-T2 binding length of 34 bases. The either of the two
35 matches at 34 bases produces the lifetime for this connectron.
The connectron 40850 has a C1-T1 binding length of 48 bases and
a C2-T2 binding length of 39 bases. The shorter of the two
matches at 39 bases produces the lifetime for this connectron.
The lifetime of this pair of anti-dominant - dominant
40 connectrons as shown in figure 10d is 1326.

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match				type					
wrm	40849	6	2	51392 CN	13819.470	13819.500	1/r g		34
ACCGAACCCAACGGCCCTCTTTAGGGCCACAAAT									
wrm	40849	6	2	51379 TN	13817.594	13817.630	1/r g		34
ACCGAACCCAACGGCCCTCTTTAGGGCCACAAAT									
wrm	40849	6	2	51392 CN	13819.470	13819.500	1/r g		34
ACCGAACCCAACGGCCCTCTTTAGGGCCACAAAT									
wrm	40849	6	2	51400 TP	13820.550	13820.583	1/r g		34
ACCGAACCCAACGGCCCTCTTTAGGGCCACAAAT									



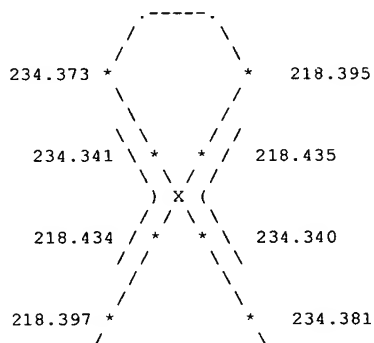
Description of an asymmetric left-right connectron pair in *H. sapiens*

The connectron 67620 has a C1-T1 binding length of 38 bases and a C2-T2 binding length of 33 bases. The shorter of the two matches at 33 bases produces the lifetime for this connectron. The connectron 67621 has a C1-T1 binding length of 41 bases and a C2-T2 binding length of 42 bases. The shorter of the two matches at 41 bases produces the lifetime for this connectron. The lifetime of this pair of dominant - anti-dominant connectrons as shown in figure 10c is 1353.

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
hsd	67620	100	1	96091 CN	1705.996	1706.033	l/r g		38
GTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC									
hsd	67620	60	1	78101 TP	218.397	218.434	l/r g		38
GTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGC									
hsd	67620	100	1	96101 CN	1705.970	1706.002	l/r g		33
AGGTCAGGAGATCGAGACCATCCTGGCTAACAC									
hsd	67620	60	1	78110 TN	234.341	234.373	l/r g		33
AGGTCAGGAGATCGAGACCATCCTGGCTAACAC									

hsd	67621	100	1	98781 CN	3142.085	3142.125	l/r g		41
CGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCC									
hsd	67621	60	1	78101 TP	218.395	218.435	l/r g		41
CGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCC									
hsd	67621	100	1	98781 CN	3142.052	3142.093	l/r g		42
GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAAC									
hsd	67621	60	1	78110 TN	234.340	234.381	l/r g		42
GAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAAC									

Found L/R DA AB - CD Connectron pair for 67620 and 67621 with a lifetime = 33 X 41 = 1353									
0.001 0.001									
218.397 218.434 234.341 234.373 --- 218.395 218.435 234.340 234.381									



**Description of an asymmetric left-right connectron pair
in *A. thaliana***

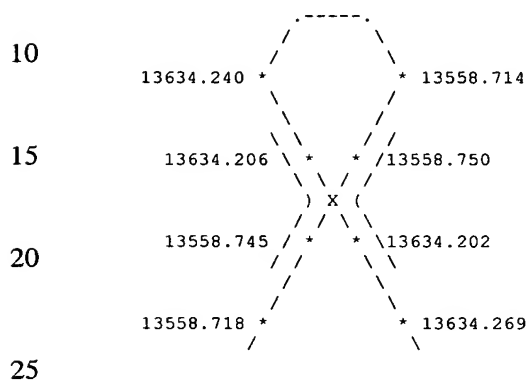
The connectron 5 has a C1-T1 binding length of 28 bases and a
5 C2-T2 binding length of 35 bases. The shorter of the two
matches at 28 bases produces the lifetime for this connectron.
The connectron 6 has a C1-T1 binding length of 37 bases and a
C2-T2 binding length of 68 bases. The shorter of the two
matches at 37 bases produces the lifetime for this connectron.
10 The lifetime of this pair of anti-dominant - dominant
connectrons as shown in figure 10d is 1036.

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match				type					
ath	5	15	3	102175 CN	12540.150	12540.174	l/r g		28
ATCATCAATGAACTCATTGGCTAAGGT									
ath	5	15	3	102902 TN	13558.720	13558.750	l/r g		28
ATCATCAATGAACTCATTGGCTAAGGT									
ath	5	15	3	102176 CN	12540.184	12540.220	l/r g		35
ACATTCATTAGTTCTGGAACGTGAATCAAGCAATG									
ath	5	15	3	103090 TP	13634.210	13634.240	l/r g		35
ACATTCATTAGTTCTGGAACGTGAATCAAGCAATG									

ath	6	15	3	103067 CP	13626.660	13626.700	l/r g		37
ATGCATCATCAATGAACTCATTGGCTAAGGTGAAGG									
ath	6	15	3	102902 TN	13558.713	13558.750	l/r g		37
ATGCATCATCAATGAACTCATTGGCTAAGGTGAAGG									
ath	6	15	3	103067 CP	13626.624	13626.691	l/r g		68
TTTAACATTCATTAGTTCTGGAACGTGAATCAAGCAATGCATCATCAATG									
AACTCATTGGCTAAGGT									
ath	6	15	3	103090 TP	13634.202	13634.270	l/r g		68
TTTAACATTCATTAGTTCTGGAACGTGAATCAAGCAATGCATCATCAATG									
AACTCATTGGCTAAGGT									

Found L/R AD AB - CD Connectron pair for 5 and 6 with a lifetime = 28 X 37 = 1036
 .005 .004

5 13558.718 13558.745 13634.206 13634.240 --- 13558.714 13558.750 13634.202 13634.269



Design of an asymmetric lower-upper connectron pair in *S. cerevisiae*

30

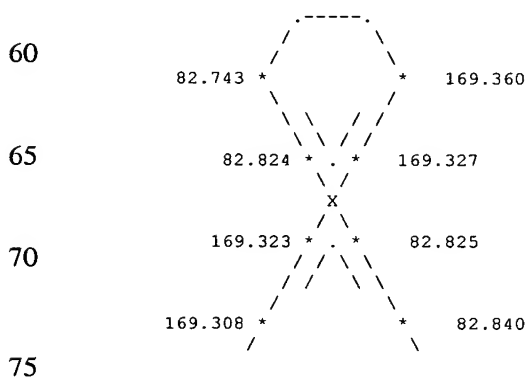
There are many ways to design a pair of connectrons. In this example we have chosen to replace the C1 source and the T1 target of the upper naturally occurring connectron with another sequence. Design of a connectron pair can be accomplished by anyone skilled the art by modifying and/or replacing any of the sources and targets in the four positions of either a lower-upper or a left-right connectron pair. A totally synthetic pair of dominant - anti-dominant connectrons could also be designed de-novo.

40

The connectron 5441 has a C1-T1 binding length of 82 bases and a C2-T2 binding length of 35 bases. The shorter of the two matches at 34 bases produces the lifetime for this connectron. The connectron 5500 has a C1-T1 binding length of 16 bases and a C2-T2 binding length of 16 bases. Either of the two matches at 16 bases produces the lifetime for this connectron. The lifetime of this pair of anti-dominant - dominant connectrons as shown in figure 9c is 544.

45

	genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
5					type					
10	sequence of match									
15	yst 5441	3	3	2944 CP	84.112	84.193	1/u g			82
	ATACGTTTGAAGAAATCACTTTATGGATTGAAACAAAGTGGAGCGAACTG									
	GTACGAACTATCAAATCATACTGATAAAAC									
20	yst 5441	3	3	2901 TP	82.743	82.824	1/u g			82
	ATACGTTTGAAGAAATCACTTTATGGATTGAAACAAAGTGGAGCGAACTG									
	GTACGAACTATCAAATCATACTGATAAAAC									
25	yst 5441	3	3	2965 CP	84.195	84.228	1/u g			34
	GAAACGTGACGGTACTCATAAAGCTAGATTTGTT									
	yst 5441	3	3	3529 TP	169.327	169.360	1/u g			34
30	GAAACGTGACGGTACTCATAAAGCTAGATTTGTT									
	yst 5500	3	3	3387 CN	151.534	151.549	1/u g			16
35	TAATTGTTGGGATTCC									
	yst 5500	3	3	3526 TN	169.308	169.323	1/u g			16
	TAATTGTTGGGATTCC									
40	yst 5500	3	3	3387 CN	151.516	151.531	1/u g			16
	AAAGGCTATAATATTA									
45	yst 5500	3	3	2905 TN	82.825	82.840	1/u g			16
	AAAGGCTATAATATTA									
50	-----									
	Found L/U DA AB - CD Connectron pair for 5441 and 5500 with a lifetime = 34 X 16 = 544									
	.001	.004								
55	82.743	82.824	169.327	169.360	---	169.308	169.323	82.825	82.840	



Design of an asymmetric lower-upper connectron pair in *H. sapiens*

There are many ways to design a pair of connectrons. In this example we have chosen to replace the C1 source and the T1

target of the right naturally occurring connectron with another sequence. Design of a connectron pair can be accomplished by anyone skilled the art by modifying and/or replacing any of the sources and targets in the four positions of either a lower-
5 upper or a left-right connectron pair. A totally synthetic pair of anti-dominant - dominant connectrons could also be designed de-novo.

The connectron 395760 has a C1-T1 binding length of 32 bases and
10 a C2-T2 binding length of 32 bases. Either of the two matches at 32 bases produces the lifetime for this connectron. The connectron 395762 has a C1-T1 binding length of 40 bases and a C2-T2 binding length of 39 bases. The shorter of the two matches at 39 bases produces the lifetime for this connectron.
15 The lifetime of this pair of anti-dominant - dominant connectrons as shown in figure 10c is 1248.

genome	Connectron id	chromosome	contig	(.groups) id	match start	match stop	type of Connectron	source of Connectron	length of match
sequence of match				type					
hsd 395760	920	19	747775	CP	17572.332	17572.363	l/r g		32
CCAGCCCTCCTCCCTCAGACCCAGGAGTCCA									
hsd 395760	922	19	765474	TN	27988.178	27988.209	l/r g		32
CCAGCCCTCCTCCCTCAGACCCAGGAGTCCA									
hsd 395760	920	19	747777	CP	17572.369	17572.400	l/r g		32
CCAGCCCTCCTCCCTCAGACCCAGGAGTCCA									
hsd 395760	922	19	765567	TP	28004.852	28004.883	l/r g		32
CCAGCCCTCCTCCCTCAGACCCAGGAGTCCA									

hsd 395762	920	19	747819	CP	17573.447	17573.486	l/r g		40
CCCCAGCCCTCCTCCCTCAGACCCAGGAGTCCAGACCCC									
hsd 395762	922	19	765474	TN	27988.176	27988.215	l/r g		40
CCCCAGCCCTCCTCCCTCAGACCCAGGAGTCCAGACCCC									
hsd 395762	920	19	747823	CP	17573.520	17573.557	l/r g		39
GGCCCCAGCCCTCCTCCCTCAGACCCAGGAGTCCAGGT									
hsd 395762	922	19	765567	TP	28004.848	28004.887	l/r g		39
GGCCCCAGCCCTCCTCCCTCAGACCCAGGAGTCCAGGT									

Found L/R AD AB - CD Connectron pair for 395760 and 395762 with a lifetime = 32 X 39 = 1248

0.006 0.004

5 27988.178 27988.209 28004.852 28004.883 --- 27988.176 27988.215 28004.848 28004.887

